STAT

[redacted] DD/CS, Memo for Chief, Eur/Afr Br, FDD
Subj: OCS Notes Relevant to Intel Methods Conference
DD/S&T 3928-66 28 July 1966

1. (Assume OCR will provide inputs re CHIVE)

2. General. Perhaps a good opening comment is that the computer itself is currently aggravating rather than alleviating the information explosion. The brute-force power of EDP has in many cases been substituted for adequate analysis of a data processing problem. This has been the case for many new computer applications. It is only after the second or third expensive iteration that a sensible perspective is obtained. We see instances where basic input data is replicated, reformatted, resorted, to the point where the output volume exceeds input by perhaps 100 times. It seems evident that the same methods applied to harnessing the information explosion elsewhere in the intelligence process must also be applied--perhaps more carefully-- in the computer environment.

3. Computer Applications. The use of the computer in intelligence processing in CIA in the past three years has expanded to the point where we see applications in almost every element of the intelligence cycle. A few examples are cited below:

a. Collection. The computer is now being used in a rudimentary fashion in collection management and in determining collection system performance. We see, for example, the "bookkeeping" functions of a requirement registry performed by the machine. More sophisticated applications are on the horizon which involve the testing of hypotheses which a manager of intelligence collection normally is concerned with, i.e., given certain collection targets and priorities and specific collection sources, how might these collection resources be allocated against these targets? At this point, however, the computer is not used as an integral part of this optimizing process; rather, it provides an answer to a specific computational translation of a collection hypothesis.

b. Data Reduction. A broad view here requires a look at areas which we consider normally under data reduction--signal analysis and the like--as well as other activities which not only filter data, but present it in a form that a "lay" analyst can use effectively. The computer has played an important role in the former area for quite some time. We see, however, demands for more sophisticated processing of higher volumes of raw data--better precision in wave form analysis, ability to perform mo e sophisticated spectral analysis and correlation, a wider variety of emitters to identify, and so on. In the latter area, we should consider topics such as automatic language translation, automatic filtering of redundant textual material, automatic dissemination, and so on. We see little operational impact of computers in these areas aside from long-term experiments with limited chance of payoff.

c. Retrieval and Analysis. Again, we can cite two subcategories: computational and non-computational data processing; and, as before, progress in the former is considerably advanced over the latter (this is independent of any value judgments on the utility of the product involved.) XXX For example, the computer has been used very effectively in filling in the complete picture of a missile or space event with only fragmentary data available. Secoddly, where a particular activity of intelligence interest can be reduced to a mathematical model, some successes have been achieved--one example here is our ability to analyze a particular air defense system.

In the area of non-computational data processing, the computer has been used principally as a file updating, sorting, searching, and printing device. Examples here include files on aircraft movements, personnel travel, facility status, and so on.

Some success has been achieved in assisting economic intelligence analysis, which is a combination of both types of data processing. Some of the first Agency applications of computers to intelligence were in this area and continue to be of some importance--computation of Soviet military expenditures an and of Soviet gross national product.

d. Production. The mechanics of providing intelligence material in a form suitable for wide dissemination involves considerable expense and in some areas is amenable to computer assistance. The two particular examples are cited here. The production of high-quality, printed material is now being assisted through the use of a computer-generated tape which is fed to an automatic photo-composing device. New editions of the National Intelligence Survey will be printed using this technique without sacrificing the aeschetic qualities of this publication (the use of multiple fonts and other complexities in print composition). Secondly, the computer is now being used on an experimental basis to provide assistance to cartographers. It is intended that a large bank of cartographic information be machine stored and selectively retrieved and plotted to provide the basic map over which specific intelligence information would be superimposed. Considerable flexibility would be available in terms of scale, projection, degree of detail, and number of geographic features shown on the map.

A basic criterion that might be applied to determine if the intelligence analyst is making effective use of the computer is,

"Would there be any disastrous effects if the computer were suddenly withheld from any of these application areas?" A candid answer in most cases, unfortunately, would be no. A basic reason for this is that the analyst has maintained separate backup facilities in many cases--perhaps because his reluctance to rely so heavily on the computer. Secondly, several of these applications are in the "why don't we try using a computer" category--without appropriate analysis of the costs, effort, and ultimate payoffs involved. There is some cause for optimism, however. Notable successes have been achieved in cases where (1) the analyst initiates the requirement (2) he understands the capabilities as well as the limitations of the computer, and (3) he permits the computer specialist to get deeply involved in the early stages of problem analysis.

4. <u>Computer Techniques</u>. Emphasis is placed here on major functional areas directly related to the computer. Some indication is given of current techniques in operational use and techniques of some promise.

a. Input. Considerable attention is now being given to the concept of "source data automation." This implies the keying of data as close to the source as possible. In some circumstances it implies direct entry of this data into the machine. A technique of interest here is optical character recognition, where the emphasis has recently shifted from an interest in reading text to the use of character readers in the place of punching paper tape or cards. Where reasonable quality control can be applied to the typewriter font, paper contrast, and character registration, available print readers are now competitive as an input device with punch card or punch paper tape readers. This input preparation technique has several obvious advantages: the use of standard office equipment for keying, decentralization of input preparation, the ability of the human to read exactly what the machine reads, ability to implement a "turn-around document" concept, i.e., computer printouts can be modified and then fed directly back as input.

Progress is also being made in design of economical secure tape typewriters. In the next few years, we expect to see such devices available for under $5,000. It is also reasonable to expect that the cathode ray tube or typewriter terminal will be used as a direct data entry device for large volume inputs. This would provide a facility for direct feedback of detected error conditions which could be serviced on the spot.

b. Processing Units. The basic design of computers has remained unchanged since the invention of the stored program concept over 15 years ago. It was designed for arithmetic operations and as such is not xx particularly appropriate to the solution of non-numeric data processing problems. However, attempts to design a machine which better fits these problems have not been very successful. For example, parallel search memories where all cells are interrogated simultaneously have not progressed beyond the laboratory. Attempts have also been made to build a machine which processes materials stored in something called an "information processing language (IPL), where data elements are stored in a tree structure and search progresses up and down branches of the structure. Another example of a special-purpose machine is one currently installed in the CIA Computer Center called the Automatic Language Processor, which employs a large-capacity disk called a "photosotre" which is used as a dictionary for lexical processing. Simulations of all of these special machines have indicated that the general purpose device which is not available is on effective and relatively inexpensive--if unsophisticated--facility for the solution of data processing problems. The principal reason for this is that the market for the general-purpose computer is extremely large and competitive, and the cost of a given amount of processing continues to decrease.

c. Files and Storage. The comments made above are applicable here as well. That is storage devices normally made available with general purpose machines cannot be considered optimal for most intelligence data processing problems. But the capacity and speed of these devices makes them economically feasible. While direct access devices such as storage drums and disks have been available for several years, our ability to use them effectively is still behind the electromechanical xe technology. For example, we find that the various segments of disk storage are really considered to be small magnetic tapes--data within each segment is usually searched serially.

d. Software. This term was coined about four years ago to identify a concept which implies that many of the facilities that must be available to users of computing equipment are appropriately implemented through a set of programs rather than through specific hardware logical features. For example, while a computer contains a control unit which automatically determines the sequence of its operations, more sophisticated automatic control features, such as automatic scheduling of major jobs on the computer, have been assigned

to programmers for implementation. Similarly, more sophisticated languages for commanding the machine--other than the basic instruction set which includes functions such as add, subtract, move, and so on--have obvious advantages to the user. These again have been implemented through programs which translate these command languages into the elemental instruction set available directly to the hardware.

This concept has mushroomed to the point where buyers of computer systems demand a wide variety of control and utility programs as an integral part of the package to be rented or bought from the manufacturer. The software concept, along with the increased capacity and speed of computing devices, has resulted in the notion of the computer as an information processing utility. As a result, we hear the term "time-sharing" quite often when we talk to computer people these days. In this kind of an environment, the computer user sitting at the console has the illusion of the complete computer being made available to him continuously. In reality, however, he is only given short bursts of the time available on a computer. These time slices might be several milliseconds for each of perhaps 100 users. This idea is similar to the method of interleaving telephone conversations in a modern communication system. If this concept is workable--and experiments thus far show promise--the computer user, whether he be a programmer or intel analyst, will have more direct access to computer facilities.

e. Analog Devices. The digital computer has for quite some time surpassed the analog computer in its speed and precision. However, an increasing amount of data which must be processed in an intelligence envirionment is collected in analog form. If the digital computer is to be used for the analysis of this raw data, more effecient means of converting the analog signal to digital form must be available. The principal limitations of conventional converters are speed, and quantitization precision. In addition, it would be desirable if the analog portion of such a converter would have feedback facilities for automatic retuning of the device to enhance the quality of the signal being processed. Converters which incorporate some of these advantages and overcome some of the speed and precision limitations previously encountered have been designed and will be used soon experimentally in the American Intel Community. Subsequent Intel Methods Conferences should look closely at the results of these experiments because of the vital role that analog collection systems promises to play in intel in the future.

5. Systems. Some basic comments shld be made here about problems of compatibility and data processing standards. We find, for example, that computer manufacturers have put us in the position where data processed on one computer is unintelligible to another. We find that programs written and tested are incomprehensible to a successor when the programmer moves on to another job. We also find cases where machine-readable products of one computer process do not match the input requirements to the next step in the process. Further, we find that people express data elements in a variety of ways, such as dates, names of people or installations, numeric quantities, and so on. The use of the computer demands a high-level of discipline at all atages of implementation and application--discipline which, up to this time, has not been exercised in the day-to-day practice of the intelligence art. The computer, of course, is not totally responsible for the attention being given to this problem now. For example, we have been plagued for a number of years with the problems of a variety of ways to transliterate cyrillic and other alphabets into the Roman alphabet and the Romanization of ideographs. Perhaps the advent of the computer will serve as a catalyst in solving some of these problems.

We should note finally that much of the blame for our failure to make progress can be placed in our failure to tackle many of our problems on a systems basis. This was alluded to in earlier remarks regarding the computer aggravating the information explosion and the need for the analyst and the computer specialist to work together. Perhaps we are merely being polite and calling these systems problems instead of candidly stating that we don't yet know how to manage our computer activities, or that we need to give more attention to that part of the intelligence process between collection and production.

STAT

/s/

Deputy Director of Computer Services

Approved For Release 2004/05/05 : CIA-RDP85B00803R000200080017-7